New Directions in DDoS Defense* (Peaceful Non-cooperation)

Guilt

You have opened my eyes, I want to travel to the peak of Himalayas and meditate for the rest of my life

Good guy

I know you are attacking machine A, here are the details of machine B

Lack of challenge

My fellow hackers will look down on me. Instead I am going to attack somebody willing to defend himself.

Why it works better 😊

Fear

Seems like a trap, he must be a genius. Instead I am going to attack somebody willing to defend himself.

Why it works better 😊

Lack of challenge

My fellow hackers will look down on me. Instead I am going to attack somebody willing to defend himself.

* from Mohit Lad (UCLA)
Outline

- Transport
- Wireless
- Other stuff

Areas

- Router interactions
  - FQ, RED → Blue, CHOKe, CSFQ, XCP…
  - Small buffer routers
- New congestion control designs
  - Delay based (Vegas)
  - Long-term TCP fair (TFRC)
  - Others: bionomial, BIC/CUBIC
- Other issues
  - Large bandwidth-delay product networks
  - Delay Tolerant Networks (DTN)
  - Congestion control outside TCP (congestion controlled UDP; general congestion management)

How does XCP Work?

Feedback = + 0.1 packet

How does XCP Work?

Feedback = - 0.3 packet

Congestion Header

Round Trip Time

Congestion Window
How does XCP Work?

Congestion Window = Congestion Window + Feedback

XCP extends ECN and CSFQ

Routers compute feedback without any per-flow state

TCP Modeling

- Given the congestion behavior of TCP can we predict what type of performance we should get?
- What are the important factors
  - Loss rate: Affects how often window is reduced
  - RTT: Affects increase rate and relates BW to window
  - RTO: Affects performance during loss recovery
  - MSS: Affects increase rate

TCP Modeling

- Given the congestion behavior of TCP can we predict what type of performance we should get?
- What are the important factors
  - Loss rate: Affects how often window is reduced
  - RTT: Affects increase rate and relates BW to window
  - RTO: Affects performance during loss recovery
  - MSS: Affects increase rate

TCP Modeling

- Given the congestion behavior of TCP can we predict what type of performance we should get?
- What are the important factors
  - Loss rate: Affects how often window is reduced
  - RTT: Affects increase rate and relates BW to window
  - RTO: Affects performance during loss recovery
  - MSS: Affects increase rate

TCP Modeling

- Given the congestion behavior of TCP can we predict what type of performance we should get?
- What are the important factors
  - Loss rate: Affects how often window is reduced
  - RTT: Affects increase rate and relates BW to window
  - RTO: Affects performance during loss recovery
  - MSS: Affects increase rate

Summary Buffered Link

- With sufficient buffering we achieve full link utilization
  - The window is always above the critical threshold
  - Buffer absorbs changes in window size
    - Buffer Size = Height of TCP Sawtooth
    - Minimum buffer size needed is 2T*C
    - This is the origin of the rule-of-thumb

Summary Buffered Link

- With sufficient buffering we achieve full link utilization
  - The window is always above the critical threshold
  - Buffer absorbs changes in window size
    - Buffer Size = Height of TCP Sawtooth
    - Minimum buffer size needed is 2T*C
    - This is the origin of the rule-of-thumb

Summary Buffered Link

- With sufficient buffering we achieve full link utilization
  - The window is always above the critical threshold
  - Buffer absorbs changes in window size
    - Buffer Size = Height of TCP Sawtooth
    - Minimum buffer size needed is 2T*C
    - This is the origin of the rule-of-thumb

Summary Buffered Link

- With sufficient buffering we achieve full link utilization
  - The window is always above the critical threshold
  - Buffer absorbs changes in window size
    - Buffer Size = Height of TCP Sawtooth
    - Minimum buffer size needed is 2T*C
    - This is the origin of the rule-of-thumb
Example

- 10Gbps linecard
  - Requires 300Mbytes of buffering.
  - Read and write 40 byte packet every 32ns.
- Memory technologies
  - DRAM: require 4 devices, but too slow.
- Problem gets harder at 40Gbps
  - Hence RLDRAM, FCRAM, etc.
- Rule-of-thumb makes sense for one flow
  - Typical backbone link has > 20,000 flows
  - Does the rule-of-thumb still hold?

If flows are synchronized

$$\sum W = \sum \frac{W}{2}$$

- Aggregate window has same dynamics
- Therefore buffer occupancy has same dynamics
- Rule-of-thumb still holds.

If flows are not synchronized

**TCP Modeling & TFRC**

- Given the congestion behavior of TCP can we predict what type of performance we should get?
- What are the important factors
  - Loss rate: Affects how often window is reduced
  - RTT: Affects increase rate and relates BW to window
  - RTO: Affects performance during loss recovery
  - MSS: Affects increase rate

$$BW = \frac{MSS}{RTT \times \sqrt{\frac{2p}{3}}}$$

- TFRC – measure each of the above and set the transmit rate to that value
Outline

- Transport
- Wireless
- Other stuff

Areas

- Ad hoc networks \rightarrow Mesh
- Sensor networks \rightarrow scale
- TCP over wireless \rightarrow DTNs & third world issues
- Mesh networks
- MAC protocols
- Software defined radios
- Chaotic wireless
- Managing wireless networks

Common Case in the Future Internet

- Historic shift from PC's to mobile computing and embedded devices...
  - >2B cell phones vs. 500M Internet-connected PC's in 2005
  - >400M cell phones with Internet capability, rising rapidly
  - Sensor deployment just starting, but some estimates ~5-10B units by 2015

Trends: Density & Management

- Densities of unlicensed devices already high
  - Limits performance due to scarce spectrum
  - Need more spectrum or more efficient use
  - Channel allocation suggests poor management
  - Makes problems such as interference worse
  - Security management probably worse
  - Won't get better \rightarrow need automation

<table>
<thead>
<tr>
<th>City</th>
<th>#APs</th>
<th>Max degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Portland</td>
<td>8883</td>
<td>54</td>
</tr>
<tr>
<td>San Diego</td>
<td>7934</td>
<td>76</td>
</tr>
<tr>
<td>San Fran</td>
<td>3037</td>
<td>85</td>
</tr>
<tr>
<td>Boston</td>
<td>2551</td>
<td>39</td>
</tr>
</tbody>
</table>

Channel

<table>
<thead>
<tr>
<th>Channel</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>51</td>
</tr>
<tr>
<td>11</td>
<td>21</td>
</tr>
<tr>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
</tr>
</tbody>
</table>
### Trends: Growing Application Diversity

**Collision Avoidance:** Car Networks

**Mesh Networks**

- **Ad-Hoc/Sensor Networks**
- **Wireless Home Multimedia**

### Trends: Spectrum Scarcity

- Densities of unlicensed devices already high
- Spectrum is scarce → will get worse
  - Improve spectrum utilization (currently 10%)

<table>
<thead>
<tr>
<th>City</th>
<th>#APs</th>
<th>Max @ 1 spot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Portland</td>
<td>8683</td>
<td>54</td>
</tr>
<tr>
<td>San Diego</td>
<td>7934</td>
<td>76</td>
</tr>
<tr>
<td>San Fran</td>
<td>3037</td>
<td>85</td>
</tr>
<tr>
<td>Boston</td>
<td>2551</td>
<td>39</td>
</tr>
</tbody>
</table>

### Implications: Spectrum Scarcity

- Interference and unpredictable behavior
  - Need better management/diagnosis tools
- Lack of isolation between deployments
  - Cross-domain and cross-technology

### Trends: Growing Deployment Diversity

- Past: largely 802.11 campus networks with laptops
- Must accommodate different optimization goals
  - Power, performance, coverage, capacity

**FUTURE**

- **Radio technology**
  - Sensor radios, 3+G cellular, Bluetooth, UWB, WiMax, software radios, and RFID
- **Deployment styles**
  - Homes, hot-spots, airports and infrastructure/municipal networks

**Devices**

Laptops, PDAs, audio/video equipment, appliances, sensors and "Constellations" of devices

**Scale**

Billions of sensors & RFID tags expected by 2015
Trends: New Technologies

- Mesh networks
  - What role should they play in access network design?
- Software radios
  - What are the right interfaces between protocol layers?
- MIMO and software-steerable antennas
- Metropolitan-area wireless networks

New Directions

- New deployments
  - FON deployment
    - Challenge: overcoming existing acceptable use policies
    - Techniques for billing, accounting, routing and discovery (of access links)
    - Municipal wireless
  - Management and diagnosis in enterprise/home/municipal networks
- Lack of privacy
  - Communications are broadcast
  - Contents of messages can be encrypted
    - Can still figure out quite a bit
- Location support
  - Place Lab shows that localization within 20 feet is possible today

New Directions

- Spectrum is scarce \(\Rightarrow\) will get worse
  - Improve spectrum utilization (currently 10%)
  - How can we manage spectrum better?
    - Dynamic re-use of spectrum?
  - Incentives/policing
    - Dynamic re-use of spectrum \(\Rightarrow\) fair allocation
      - Different optimization goals \(\Rightarrow\) difficult to reason about behavior
      - Need to monitor spectrum to ensure fair-play
      - How do we police spectrum usage?
  - 802.11 MAC… a disaster?
    - How are wireless networks really used and what is a reasonable MAC design?

Outline

- Transport
- Wireless
- Other stuff
Distributed Multiplayer Game Server

- Well connected game servers located throughout USA
- Clients connect to nearest game server
- Good interactive response
- Servers communicate using Mercury

What do Multiplayer Games Look Like?

- Large shared world
  - Composed of map information, textures, etc
  - Populated by active entities: user avatars, computer AI’s, etc
  - Only parts of world relevant to particular user/players

Demo

Colyseus Architecture Overview

1. Specify Predicted Interests: (5 ≤ x ≤ 65 & 10 ≤ y ≤ 200) TTL 30sec
2. Locate Remote Objects: P3 on s2, P4 on s2
3. Register Replicas: R3 (to s2), R4 (to s2)
4. Synchronize Replicas: R3, R4
5. Infer Interests: R3, R4, P2
6. Optimize Placement: migrate P1 to server s2

Game World

Player 1

Player 2

Object Location

Replica Management

Object Placement

Object Store

server s1

server s2

Mercury

P1

P2

R1

R2

R3

R4

R5

P3

P4
**Player Update – Solution**

- Use publish-subscribe to "register" long-lived distributed lookups
- Publications created each time object is updated (or periodically when no update is done)
  - Publication contents = state of object

```
(x ≥ 50
x ≤ 150
y ≥ 150
y ≤ 250
```

**MERCURY [Sigcomm 2004]**

- Send subscription to any one attribute hub
- Send publications to all attribute hubs
- Tunable number of long links → can range from full-mesh to DHT-like

**Distributed Hash Tables (DHT)**

```
x = 1
```

**Using DHTs for Range Queries**

- No cryptographic hashing for key → identifier

```
Query: 6 ≤ x ≤ 13
key = 6 → 0xab
key = 7 → 0xda
key = 13 → 0x12
Query: 6 ≤ x ≤ 13
```
10

Using DHTs for Range Queries

- Nodes in popular regions can be overloaded
- Load imbalance!

DHTs with Load Balancing

- Mercury load balancing strategy
  - Re-adjust responsibilities

DHTs with Load Balancing

- Range ownerships are skewed!

Ideal Link Structure

- Each routing hop may not reduce node-space by half!
  - $\Rightarrow$ no log(n) hop guarantee
Internet Bottlenecks

- Where are bottlenecks in the Internet?
  - Ignoring access links
- What is the capacity of these bottlenecks?
- Initial results
  - There is a lot of available bandwidth in the Internet today
    - > 45Mbps on 50% of paths!

How can stub networks like CMU route traffic around bottlenecks?
- Using multiple ISPs can…
  - Improve performance and reliability of Internet connectivity
  - Make Internet routing robust to failures and attacks
- But need…
  - Techniques for stub domains to choose providers
  - Monitoring tools to track changes in Internet performance
  - Dynamic control over chosen routes

The effective use of multiple ISPs (multihoming) by stub networks

How will bottlenecks change over time?
- Analyzed the combination of Internet topology and routing
- Key results
  - Congestion scales poorly in Internet-like graphs
  - Policy-routing does not worsen the congestion
  - Alleviation possible via simple, straightforward mechanisms

Uniformly scale all capacities?
- Scale some links faster?
- Moore’s-law like scaling sufficient?